

Anhang

DGAI-Reanimationsregister 2014 - 2021

Fragestellung an den Statistiker

1. Ist es möglich, dass unter Verwendung der Daten aus der gleichen Quelle, allerdings über einen unterschiedlichen Zeitraum, eine Zunahme der Laien-Reanimation um 37 % bei Abnahme der 24 h Überlebenden um 21 % und der Lebend-Entlassenen um 24 % (s. Abbildung DGAI 2014-2020) gezeigt wird, hingegen bei Anwendung durchaus anspruchsvoller statistischer Verfahren, bei deutlich zunehmender Laienreanimationsrate eine über die Jahre ansteigende Krankenhausentlassungsrate festgestellt werden konnte (s. Hubar et al. 2006-2020)?
2. In der Publikation Hubar et al. 2022 wurde eine „Datenbereinigung“ vorgenommen: Für den Zeitraum 2016-2020 werden von der DGAI hierfür 23.404 Patientendaten genannt, in der obigen Graphik (nicht dargestellt) nur 28.105. Eine „Datenbereinigung“ von immerhin 17 %?

Zitate aus der Stellungnahme des Statistikers #6887

Diese Stellungnahme bezieht sich auf die beiden Fragestellungen (inclusive erhaltener Zusammenfassung) sowie Hubar et al. (2022) im Folgenden „Paper“.

Die Diskrepanz zwischen den Ergebnissen lässt sich schlicht über die verwendete Methodik erklären. Dafür werden zwei Aspekte berücksichtigt: Datenbereinigung und die verwendeten statistische Verfahren.

Datenbereinigung

Die Datenbereinigung bezieht sich darauf, dass die Daten der vorhandenen Quelle genauer untersucht werden und mögliche Ausreißer und (unsinnige) Datenpunkte ausgeschlossen werden. Ein unsinniger Datenpunkt bezieht sich auf Angaben, die keinen Sinn ergeben und/oder fehlerhaft sind (z.B. ein Alter von 253). Daneben gibt es als Teil der Datenbereinigung das Filtern der Daten. Werden die reinen Raten von Laien-CPR, Überlebensrate und lebend entlassen-Rate betrachtet, so wird hier nicht berücksichtigt, dass z.B. eine 90-jährige Person zwar wiederbelebt werden konnte, aber diese kurz danach an z.B. Nierenversagen verstorben ist. Ggf. wird an einer solchen Stelle die Entscheidung getroffen, z.B. nur Personen zu berücksichtigen, deren Todesursache nicht unmittelbar mit dem Grund für CPR

zusammenhängt. Inwiefern dies hier Anwendung fand, ist nicht ersichtlich, aber um einen tatsächlichen Zusammenhang zu beschreiben unumgänglich.

Das Paper beschreibt, dass alle Patienten mit unvollständigen Angaben ausgeschlossen wurden. Daneben wurde für die Grafik der Zusammenfassung nur ein Ausschnitt der Daten gewählt. Bei zeitlichen Entwicklungen kann dies eine große Rolle spielen. Das Paper hat verschiedene Zeitfenster für die Analyse gewählt, weil in diesen jeweils andere Richtlinien für CPR gelten. Ebenso betrachtet das Paper die lebend-entlassen-Rate genauer, hier wird unterschieden wie der neurologische Zustand des Patienten ist. Dies ist wiederum nicht direkt vergleichbar, weil die ursprüngliche Rate aufgespalten wird.

Bei der Grafik (DGAI 2014-2021) wurde nichts zur Datenbereinigung geschrieben, weshalb zu vermuten ist, dass dies die Rohdaten sind, während im Paper eine Datenbereinigung durchgeführt wurde. Dadurch ist hier die Datenlänge nicht mehr dieselbe und nur noch bedingt vergleichbar.

[Kommentar RZ: Man darf unterstellen, dass die DGAI nur bereinigte Daten publiziert, Zitat aus DGAI 2021: Die Gesamtdaten des Deutschen Reanimationsregisters zwischen dem 01.01.2020 und dem 31.12.2020 sowie die Daten einer Referenzgruppe von 27 teilnehmenden Notarzt- und Rettungsdiensten aus demselben Zeitraum bilden die Grundlage dieses Berichts. Die Referenzgruppe setzt sich aus solchen Standorten zusammen, deren Daten die folgenden Einschlusskriterien und damit die Kriterien für eine besonders hohe Datenqualität erfüllen:

- Inzidenz für Reanimationen > 30/100.000 Einwohner und Jahr
- Jemals ROSC (Return of Spontaneous Circulation) < 80 %
- RACA (ROSC after Cardiac Arrest)-Score berechenbar > 60 %
- Anteil an dokumentierten Weiterversorgungen von mindestens 30 %.]

17% der Daten, die wegfallen, ist nicht unüblich. Die Zahlen sind aber nicht zu vergleichen, weil die Graphik die Daten von 2014-2021 erfasst. Damit überschneiden sich dort zwei Zeitfenster des Papers plus das Jahr 2021.

[Kommentar RZ: Es werden gleiche Zeiträume verglichen, nämlich 2016-2020 (DGAI und Paper)].

Als Richtwert für (unproblematische) Datenbereinigung können bis 25% genommen werden, denn jede Datenbank ist nur so gut, wie die Person, die sie pflegt. Dadurch kommt es nicht selten bei Datenerfassungen über viele Jahre hinweg zu Problemen, die leider nur durch Ausschluss umgangen werden können. Bei allem was darüber liegt wird, wenn das Paper statistisch gut aufgestellt ist, kurz erklärt, warum so viele Datenpunkte ausgeschlossen werden mussten.

Statistische Verfahren

Der wichtigere Punkt sind hier die statistischen Verfahren. Der Grund für die Anwendung aufwändiger statistischer Verfahren ist, dass einfache Verfahren (z.B. Korrelation) nur einfache

Zusammenhänge beschreiben können. Bei komplexeren Zusammenhängen (wie bei der Überlebensrate von Patienten nach CPR zu erwarten) müssen dementsprechend andere Verfahren gewählt werden, die den tatsächlichen Zusammenhang besser abbilden können. In dem Paper wurde u.a. eine Regressionsanalyse durchgeführt. Diese Analysen haben den großen Vorteil, dass sie den Einfluss anderer Variablen auf dieselbe zu erklärende Variable berücksichtigen. Bei dieser Regression wurde als zu erklärende Variable das Überleben des Patienten mit guten neurologischen Behandlungsergebnis gewählt (also Ja, überlebt oder Nein, verstorben; daher eine logistische Regression). Ob ein Patient überlebt, hängt zu jedem Zeitpunkt von einer Vielzahl von Einflüssen ab, was bereits intuitiv zu erwarten ist. Das Paper untersucht diese Einflüsse und konnte ein Modell aufstellen, dass die Entlassungsrate mit guten neurologischen Behandlungsergebnis moderat erklärt (Nagelkerkes R^2 von 0,43 wird als moderat interpretiert).

Das zweite Verfahren ist die sog. Matching-Pair Analyse. Diese untersucht ob es von dem ersten Zeitfenster, im Vergleich zum dritten Zeitfenster statistisch signifikante Unterschiede in Variablen vorliegen. Dies bedeutet das zwei verschiedenen Perioden von verschiedenen CPR-Richtlinien verglichen werden, um zu überprüfen, ob die Rate der Outcome Variablen sich signifikant verändert hat. Diese Ergebnisse sind wiederum nicht vergleichbar mit der Grafik (DGAI 2014-2021), da hier nur ein Zeitfenster betrachtet wird.

Im Vergleich dazu wird in der Grafik nur deskriptiv, also beschreibend, die Entwicklung betrachtet. D.h. hier kann zwar gesagt werden, dass eine bestimmte Rate sinkt oder steigt, aber nicht, ob das statistisch signifikant ist und ein Trend ist, der sich so fortführen wird oder nicht. In der Conclusion wird von Proportionen gesprochen, d.h. es werden nicht die absoluten Zahlen betrachtet, sondern wie viel Prozent des Gesamtgutes eine Eigenschaft haben. Diese prozentualen Beträge werden dann verglichen. Die Stichprobengröße geht dabei in die Berechnung des Konfidenzintervalls ein, dieses beschreibt dann, ob ein Verhältnis zwischen zwei prozentualen Beträgen signifikant ist. D.h. wenn das Konfidenzintervall die Null nicht miteingeschlossen hat, gibt es ein statistisch signifikantes Verhältnis zwischen den zwei Beträgen. Dazu würden Matching-Pair gebildet, d.h. für die Patienten der ersten Stichprobe wurden statistische Zwillinge in der dritten Stichprobe gesucht und nur diese miteinander verglichen. In welchen Variablen diese Zwillinge identisch sein mussten ist der Tabelle 2 zu entnehmen.

Statistisch ist diese Aussage des Papers völlig zulässig, weil sie nun mal die Größe der Stichproben mitberücksichtigt. Das genaue Matching Verfahren wird im Paper nicht erwähnt, lediglich die verwendete Software. Über Matching Verfahren gibt unter Statistiker:innen allerdings viel Diskussion, ob diese wirklich gut sind oder nicht. D.h. dies ist reine Meinungssache. Der Unterzeichner erachtet sie als sinnvoll, wenn sie gut durchgeführt werden und das Matching sehr gut durchdacht ist.

Die Zeiträume an sich zu vergleichen ist mehr inhaltlicher als statistischer Natur. Der Unterzeichner vermutet, dass vom 2. zum 3. Zeitraum kein großer Unterschied festzustellen war, genau wie vom 1. zum 2. Zeitraum. Der Grund dafür kann schlicht sein, dass es eine längere Zeit braucht, bis neue Richtlinien in der Bevölkerung und im medizinischen Personal wirklich angekommen und angewendet wird. Aber wie gesagt, dies ist nur eine Vermutung (dem Unterzeichner ist erst im Zuge dieser Stellungnahme klar geworden, dass sich die CPR-Richtlinien geändert haben). Warum tatsächlich nur von dem Vergleich zwischen 1. und 3. Zeitraum berichtet wurde, kann nur von den Autor:innen beantwortet werden.

Die Belege sind Table 3 – Multivariate logistical analysis zu entnehmen. Diese Tabelle beschreibt die Ergebnisse der logistischen Regression mit CPC 1,2 als abhängige Variable (eine dichotome Variable mit ja oder nein). Jede signifikante Variable (p -Wert $< 0,05$) hat einen statistisch signifikanten Einfluss darauf, ob ein Patient CPC 1,2 hat oder nicht. Um den Einfluss genauer zu bestimmen wird der Regression coefficient B betrachtet, ist er positiv so ist der Einfluss positiv, analog für negativ. Für Bystander CPR ist B 0,25 mit einem p -Wert von $< 0,001$, d.h. Patienten, die Bystander CPR erhalten haben, haben auch eher CPC 1,2, im Vergleich zu Patienten, die kein Bystander CPR erhalten haben. $\text{Exp}(B)$ gibt dann die Odds Ratio (= Chancenverhältnis) an. Hierbei heißt ein $\text{Exp}(B)$ von 1,28, dass Patienten, die Bystander CPR erhalten haben eine 1,28-Fache Chance haben ein CPC 1,2 zu haben, im Vergleich zu jenen, die kein Bystander CPR erhalten haben.

Von associated wird danach gesprochen, weil das Modell nur moderat ist und nicht sehr gut. D.h. im Paper konnte gezeigt werden, dass es dort positive Zusammenhänge gibt, aber wie groß diese genau sind müsste in einem anderen Modell genauer untersucht werden. Oder mit Variablen, die in der Datenbank nicht hinterlegt sind.

Zusammenfassung

Die Daten kommen zwar aus derselben Quelle, jedoch mit verschiedenen zeitlichen Abschnitten, in denen jeweils andere Richtlinien galten. Dazu kommt die Datenbereinigung der Daten für das Paper, diese lassen sich nur bedingt mit den Rohdaten vergleichen. [fraglich, s.o.] Die verwendeten statistischen Verfahren sind es aber im Endeffekt, die diese Ergebnisse nicht miteinander vergleichbar machen. Der Grund liegt darin, dass die Verfahren des Papers mehr Einflüsse berücksichtigen und somit sehr wahrscheinlich näher die Realität abbilden, während die Abbildung (DGAI 2014-2021) die Daten nur deskriptiv beschreibt. D.h. die Grafik zeigt (im Sinne von beweist) nichts, sie zeigt nur den Verlauf der Raten dieser Jahre. Für tatsächliche Aussagen zu treffen, müssen immer statistische Verfahren angewendet werden.

Januar 2023

Dr. Robert Grünwald
Geschäftsführer
Novostat GmbH, 8832 Wollerau, Schweiz.